Research Article

THINKING OF CRITICAL WORDS DURING STUDY IS UNNECESSARY FOR FALSE MEMORY IN THE DEESE, ROEDIGER, AND MCDERMOTT PROCEDURE

John G. Seamon, Ihno A. Lee, Sarah K. Toner, Rachel H. Wheeler, Madeleine S. Goodkind, and Antoine D. Birch

Wesleyan University

Abstract—*Do participants in the Deese, Roediger, and McDermott (DRM) procedure demonstrate false memory because they think of nonpresented critical words during study and confuse them with words that were actually presented? In two experiments, 160 participants studied eight visually presented DRM lists at a rate of 2 s or 5 s per word. Half of the participants rehearsed silently; the other half rehearsed overtly. Following study, the participants' memory for the lists was tested by recall or recognition. Typical false memory results were obtained for both memory measures. More important, two new results were observed. First, a large majority of the overt-rehearsal participants spontaneously rehearsed approximately half of the critical words during study. Second, critical-word rehearsal at study enhanced subsequent false recall, but it had no effect on false recognition or remember judgments for falsely recognized critical words. Thinking of critical words during study was unnecessary for producing false memory.*

People can misremember in two general ways. They can make errors of *omission* in which they fail to recollect a previous event, and they can make errors of *commission* in which they recollect a prior event differently from how it was originally experienced. Psychologists have had a long interest in studying omission errors because these errors are so commonly experienced, but the study of commission errors has produced a more uneven record because commission errors are less subjectively obvious. People know when they fail to recollect, but they are generally oblivious to their failure to recollect accurately. Only recently has the study of commission errors, more commonly known as *false memories*, received intense experimental investigation. One procedure for studying false memory was outlined by Roediger and McDermott (1995), following earlier research by Deese (1959).

Deese (1959) presented participants with lists of semantic associates that converged on nonstudied words. For example, for the list consisting of *thread*, *pin*, *eye*, *sewing*, *sharp*, *point*, *pricked*, *thimble*, *haystack*, *pain*, *hurt*, and *injection*, the converging associate was *needle*. Following study of lists of this type, people often erroneously recalled the nonstudied converging associates, hereafter referred to as *critical words*. Roediger and McDermott (1995) revived interest in this procedure by demonstrating that it can be extended to the study of false recognition and metamemory judgments. Numerous studies have since shown that this Deese, Roediger, and McDermott (DRM) procedure readily produces high levels of false recall or recognition for critical words. Yet, despite intense experimental study, no consensus exists on the explanation of this memory illusion.

Address correspondence to John Seamon, Department of Psychology, Wesleyan University, Middletown, CT 06459; e-mail: jseamon@wesleyan.edu.

The different explanations of the DRM illusion include Underwood's (1965) *implicit-activation response hypothesis*, Roediger and his colleagues' (Roediger, Balota, & Watson, 2001; Roediger & McDermott, 2000) *activation-monitoring approach*, and Brainerd and Reyna's (1996, 1998) *fuzzy-trace theory*. According to Underwood, when participants encode words at study, they also activate representations for semantic associates of those words. This activation is based on a conscious process of covert verbalization wherein semantically related words are articulated, but not overtly spoken (Underwood, 1965, p. 122). On a subsequent memory test, the nonstudied, but covertly verbalized words lead to false recognition because, from the participants' perspective, they represent stimuli that actually occurred.

The activation-monitoring approach, put forth by Roediger and his colleagues, is an extension and modification of the activation hypothesis. Based, in part, on research that shows that false memory is still observed in the DRM procedure following study exposures of only 20 ms per word (Roediger, Balota, & Robinson, 2002; Seamon, Luo, & Gallo, 1998), Roediger et al. (2001) suggested that critical words can be activated either consciously by elaborative processing or automatically through the spread of activation within a semantic network, as when list words are presented too quickly for conscious processing.

According to Brainerd and Reyna's fuzzy-trace theory, memory judgments are based on verbatim or gist traces that are established by parallel processes at the time of study. Verbatim traces represent the surface details of physical stimuli, whereas gist traces represent the meaning or theme of the stimuli. Accurate recall of studied words in the DRM procedure is driven largely by verbatim traces, whereas false recall of critical words is based predominantly on gist traces (Brainerd & Reyna, 1996, 1998; Brainerd, Wright, Reyna, & Mojardin, 2001). Because false memory is still observed under conditions that yield poor memory for list words (e.g., Seamon et al., 1998), gist traces could be established at an earlier stage of processing than verbatim traces (Brainerd et al., 2001).

For the present research, the important point is that the activation-monitoring approach and fuzzy-trace theory are alike in assuming that false memory in the DRM procedure can result from a fast-acting, largely automatic activation or generation process, whereas Underwood's activation hypothesis holds that false memory is the result of the conscious activation of critical words during study. In this experiment, we tested these assumptions directly by monitoring the thoughts of participants as they studied DRM lists. Using an overt-rehearsal procedure originally devised by Rundus (1971), we sought to answer two fundamental questions about the DRM procedure. First, will participants spontaneously rehearse nonpresented critical words during study? Second, if so, what are the memorial consequences of critical-word rehearsal at study on subsequent measures of recall, recognition, and remember/know judgments in recognition?

Two studies provide partial answers. First, Mather, Henkel, and Johnson (1997) used a memory-characteristics questionnaire with the DRM procedure and reported that participants recalled thinking about critical words during study while trying to remember list words. Second, Goodwin, Meissner, and Ericsson (2001) used overt rehearsal with the DRM procedure and found that rehearsal of critical words was moderately correlated with subsequent false recall. Whereas both experiments indicate that critical words are thought of during study, neither experiment provided data on the proportion or frequency of critical-word rehearsal and, more important, whether prior rehearsal had any causative effect on subsequent memory performance. With this study, we sought to determine that relationship.

METHOD

Participants

The participants were 160 Wesleyan University students who received credit in an introductory psychology course or served as paid volunteers. None had taken part in any related research.

Materials

We used 16 DRM word lists obtained from Roediger and McDermott (1995). Each list was composed of 15 words that were all converging associates of a nonpresented critical word. Within each list, the order of the words was constant, and the strongest associates to the critical word normally occurred first. The stimulus lists were divided into two sets of 8 lists, labeled A and B, for counterbalancing. Within each condition, half of the participants received Set A for study and half received Set B. The stimulus set that was not used during study provided distractors for participants tested by recognition.

Procedure

Intentional-learning instructions were used. Participants looked at the screen of an Apple computer, where lists of words were presented, and were instructed to remember as many words as possible for a memory test that would follow the last list. Half of the participants were told that their memory would be tested by recall, and half were told that it would be tested by recognition. Following a practice set of unrelated words and a written recall or recognition test, all participants were presented with eight lists of 15 words in blocked fashion. The words were typed in 24-point, uppercase, black Helvetica font, and they were presented sequentially in the center of the screen at a constant rate of either 2 s or 5 s per word with no delay between words. These presentation rates were selected because the 2-s rate is typical for DRM research (e.g., Roediger & McDermott, 1995; Seamon et al., 2002; Toglia, Neuschatz, & Goodwin, 1999), whereas Rundus (1971) used a rate of 5 s for overt rehearsal. Successive lists were separated by a red, 1-s "Next List" prompt, and participants were not informed about the semantic relatedness of the words in the study lists.

While the word lists were shown during study, half of the participants in each presentation-rate condition engaged in silent rehearsal, and half used overt rehearsal. The silent-rehearsal condition was the same as that used by previous researchers with the DRM procedure, whereas the overt-rehearsal condition was new. In adapting Rundus's overt-rehearsal procedure, we instructed participants to think out loud by saying any word or thought that came to mind, even if it was unrelated to the word lists, while the words were presented on the screen. These participants were asked to speak continuously in a conversational tone for the entire duration of the study words. Following the instructions and practice, these participants were left alone in a sound-controlled room so that the presence of the experimenter might not inhibit their speech, which was recorded by a tape recorder placed next to the computer.

After the study lists were presented, half of the participants in each condition were given a free-recall test in which they were asked to write down as many of the previously shown words as they remembered. No time limit was placed on the test. The other half of the participants received a written recognition test that consisted of 64 words: 3 words from each studied list (Serial Positions 1, 8, and 10), the nonpresented critical word from each studied list (*related critical words*), 3 words from each nonstudied list (Serial Positions 1, 8, and 10), and the critical word from each nonstudied list (*unrelated critical words*), all presented sequentially on a test sheet in a random order. Each word was accompanied by the words "Yes" and "No" for a yes/no recognition decision and the letters "R" and "K" for a remember/know judgment. The participants were asked to examine the words in sequential order and make a recognition decision for each word. They were told to circle "Yes" for any word that they recognized from the study lists and "No" for any word they failed to recognize. In addition, they were instructed to make a remember or know judgment for each recognized word by circling either "R" or "K" on the answer sheet. They were told to circle "R" if they consciously recalled the word from the study lists and "K" if they felt sure that the word was presented, but they could not remember its specific occurrence. Following testing, all participants were debriefed.

RESULTS AND DISCUSSION

Four sets of analyses are reported. First, from the rehearsal tapes of the overt-rehearsal participants, we calculated the mean proportions and frequency of rehearsal for list words and critical words during study. Second, we obtained the mean proportions of accurate recall of list words and false recall of critical words. Third, we measured correct and false recognition, as well as sensitivity and response bias, for list words and critical words. Finally, for the overt-rehearsal participants, we conditionalized false memory performance in terms of the presence or absence of critical-word rehearsal to observe any effect of rehearsal on false recall, recognition, and remember judgments. Because list set did not systematically affect performance for any dependent measure, this variable is not considered further.

Rehearsal During Study

Table 1 presents the mean proportions and frequency of rehearsal for list words and critical words at each presentation rate. Because the 80 overt-rehearsal participants were all exposed to the same rehearsal condition during study, regardless of whether they were subsequently tested by recall or recognition, their results were pooled, according to presentation rate at study. The results indicate that all participants rehearsed at least one list word, and a large majority of the participants at each presentation rate also rehearsed at least one critical word. Slowing the presentation rate from 2 s to 5 s increased the proportion of participants who rehearsed at least one critical word, $z = -1.95$, $p < .05$ (one-tailed).

In addition, the participants rehearsed most list words, especially at the slower rate, and they rehearsed approximately half of the critical words at both rates. The results of separate analyses of variance with

Table 1. *Mean proportions and average frequency of rehearsal per word for list words and critical words*

Presentation rate and word type	Rehearsal: Proportion of participants	Proportion of words	Frequency per word
2 s			
List words	1.00	.79	1.00
Critical words	.73	.45	1.25
5 s			
List words	1.00	.98	3.47
Critical words	.90	.59	2.20

Note. The first column shows the proportion of 40 participants who rehearsed at least one word in each condition, the second column shows the average number of words rehearsed by all participants in each condition, and the third column shows the average frequency of rehearsal by all participants in each condition.

presentation rate and list set as between-participants variables indicated an effect of rate on the proportion of words rehearsed, for both list words, $F(1, 76) = 14.28$, $MSE = 0.73$, $p < .001$, and critical words, $F(1, 76) = 3.56$, $MSE = 0.40$, $p < .07$. Slowing the presentation rate allowed for more list words and critical words to be rehearsed. Essentially the same effect of presentation rate was observed on the mean frequency of rehearsal. Participants rehearsed list words and critical words approximately once each on average with the 2-s rate, but two to three times each with the 5-s rate. This increase in rehearsal frequency with presentation rate was reliable for both list words, $F(1, 76) = 95.78$, $MSE = 122.51$, $p < .0001$, and critical words, $F(1, 76) = 5.60$, $MSE = 18.08$, $p < .05$.

In summary, as indicated in Table 1, more participants rehearsed list words than critical words, participants rehearsed a greater proportion of list words than critical words, and, at least at the 5-s rate, they rehearsed list words more frequently than critical words. Yet critical words were rehearsed, and presentation rate influenced rehearsal of list words and critical words in a similar manner. The answer to our first research question is an unambiguous "yes": Participants will rehearse nonpresented critical words during study. A large majority of the participants spontaneously rehearsed approximately half of the critical words.

Accurate and False Recall

Table 2 shows the mean recall proportions for list words and critical words. Three important observations can be made. First, the recall results demonstrate a strong false memory effect. Second, they show that overt rehearsal at study did not alter performance, as the mean recall proportions for list words and critical words were similar for the silent- and overt-rehearsal groups. Third, slowing the presentation rate from 2 s to 5 s per word enhanced accurate recall for both groups, but not false recall. These results were supported by analyses of variance that showed an effect of rate on the accurate recall of list words, $F(1, 72) = 27.97$, $MSE = 0.29$, $p < .0001$, but not on the false recall of critical words, $p > .20$. Although McDermott and Watson (2001) have already observed a dissociative effect of presentation rate on accurate and false recall, the present finding is noteworthy because lengthening the presentation rate from 2 s to 5 s increased both the proportion and the frequency of critical words rehearsed, but did not increase false recall. Differences between rehearsal groups and rate-by-group interactions were not observed, all Fs < 1.0. Finally, the mean frequency of recall intrusions was low for each rate in both rehearsal conditions (silent: 0.95, 1.95; overt: 1.50, 1.15), which did not differ from each other, all ps $> .10$.

Accurate and False Recognition

Table 3 shows the mean proportions of recognized words for each presentation rate, rehearsal condition, and word type. Accurate memory for list words and false memory for related critical words are indicated by greater recognition proportions for studied than nonstudied words and for related than unrelated critical words in all conditions. In addition, list words and related critical words typically yielded more remember than know judgments for recognized words, whereas the reverse was true for the nonstudied and unrelated distractors. Because the overt-rehearsal condition produced lower baseline false alarm rates than the silent-rehearsal condition, all data were corrected by a high threshold measure and analyzed in terms of sensitivity and response bias.

True Memory Sensitivity and False Memory Susceptibility

The estimate of sensitivity, P_r, is defined as the number of hits (H) minus the number of false alarms (FA), with higher scores indicating greater sensitivity than lower scores. The estimate of response bias, B_r,

Table 2. *Mean recall proportions for list words and critical words*

Presentation rate and word type	Rehearsal condition: Silent	Overt
2s		
List words	.27	.28
Critical words	.30	.33
5 s		
List words	.41	.38
Critical words	.28	.23

is defined as $FA/[1 - (H - FA)]$. Bias values of .5 indicate a neutral bias, values greater than .5 indicate a liberal bias, and values less than .5 indicate a conservative bias (Snodgrass & Corwin, 1988). Two sets of sensitivity and bias measures were computed to provide separate estimates of item-specific memory and false memory susceptibility. For item-specific memory, the mean estimates of sensitivity and bias were computed from hits and false alarms to studied and nonstudied list words. For false memory susceptibility, these estimates were computed from false alarms to related and unrelated critical words. This procedure treated the false recognition of related critical words as "hits" to measure the extent to which participants were fooled by these words (cf. Koutstaal & Schacter, 1997; Seamon, Luo, Schlegel, Greene, & Goldenberg, 2000; Seamon, Luo, Shulman, Toner, & Caglar, in press).

The results, shown in Table 3, indicate that sensitivity scores for list words were greater for overt than silent rehearsal, $F(1, 72) = 29.85$, $MSE = 0.51$, $p < .0001$, and for the 5-s than the 2-s rate, $F(1, 72) = 23.38$, $MSE = 0.45$, $p < .0001$, but neither of these variables influenced sensitivity scores for critical words, both *p*s $> .10$. Finally, a general conservative response bias was observed in most conditions (B_rs $< .5$). Thus, neither the rehearsal condition nor presentation rate influenced the recognition of critical words, just as these variables did not influence false recall.

Rehearsal and False Memory

Two different analyses were conducted to determine the effect of rehearsal on subsequent false memory. First, the overt-rehearsal participants were divided into four groups of 20, depending on whether they saw the word lists at a 2-s or 5-s rate and whether they received a recall or recognition test. From the rehearsal tapes, we tabulated how often each critical word was rehearsed (0 times, 1 time, etc.) and whether it was subsequently recalled, recognized, or judged as remembered if recognized. Collapsing across participants in each condition (20 participants × 8 critical words) yielded a maximum of 160 observations that were used to generate the conditional probabilities of false recall, recognition, and remember judgments as a function of rehearsal frequency during study. For recall and recognition, the number of observations was less than 160 in some conditions because a few participants at the start of the experiment erroneously received one of the critical words as a list word. For remember judgments, the number of observations was less than 160 because this judgment was always dependent on recognition.

Table 4 presents these conditional probabilities. For recall, the conditional probability was greater if a critical word was rehearsed than if it was not, for both presentation rates. Because the number of observations was small for some of the individual rehearsal frequencies, we collapsed the results to compare the conditional probabilities for critical words that were never rehearsed (column 1) and those that were rehearsed one or more times (column 6). One-tailed *z*-test comparisons of these proportions indicated that rehearsing the critical words one or more times led to more false recall than no rehearsal of these words, both at the 2-s rate, $z = 1.84$, $p < .05$, and at the 5-s rate, $z = 3.52$, $p < .001$. Rehearsal of critical words during study was not essential to false recall, but clearly enhanced it.

Table 3. *Mean recognition proportions, remember/know judgments, and sensitivity (P_r) and bias (B_r) measures for list words and critical words*

	Rehearsal condition			
	Silent		Overt	
Presentation rate and word type	Recognition	(R/K)	Recognition	(R/K)
2 s				
List words				
Studied	.75	(.56/.19)	.83	(.56/.26)
Nonstudied	.16	(.05/.11)	.11	(.01/.10)
P_r (B_r)	.58 (.37)		.71 (.37)	
Critical words				
Related	.74	(.49/.24)	.73	(.44/.28)
Unrelated	.32	(.09/.23)	.19	(.03/.16)
P_r (B_r)	.46 (.61)		.53 (.39)	
5 s				
List words				
Studied	.82	(.56/.26)	.93	(.75/.18)
Nonstudied	.13	(.05/.08)	.03	(.01/.02)
P_r (B_r)	.69 (.36)		.90 (.16)	
Critical words				
Related	.76	(.49/.27)	.60	(.29/.30)
Unrelated	.16	(.04/.11)	.04	(.01/.03)
P_r (B_r)	.61 (.43)		.56 (.10)	

Note. R = remember judgment; K = know judgment. Instances in which remember and know proportions do not sum to the overall proportion reflect rounding to two decimal places.

Table 4. *Mean conditional probabilities of recall, recognition, and remember judgments for critical words*

Memory measure and presentation rate	Rehearsal frequency during study					
	0 times	1 time	2 times	3 times	4 or more times	Mean of 1 or more times
Recall						
2 s (n = 157)	.27 (94)	.48 (23)	.36 (11)	.39 (18)	.36 (11)	.41 (63)
5 s (n = 158)	.12 (86)	.21 (19)	.25 (15)	.40 (10)	.54 (28)	.37 (72)
Recognition						
2 s (n = 160)	.74 (81)	.57 (30)	.53 (19)	.77 (13)	.82 (17)	.65 (79)
5 s (n = 150)	.60 (40)	.50 (28)	.65 (20)	.60 (20)	.60 (42)	.58 (110)
Remember						
2 s (n = 111)	.62 (60)	.71 (17)	.60 (10)	.30 (10)	.64 (14)	.59 (51)
5 s (n = 88)	.38 (24)	.64 (14)	.38 (13)	.50 (12)	.56 (25)	.53 (64)

Note. The conditional probability means are based on 40 participants tested by recall and 40 participants tested by recognition. The number of observations for each probability is shown in parentheses. The final column shows weighted means for items rehearsed one or more times.

A different pattern of results was observed for the conditional probabilities for recognition and remember judgments. Rehearsing a critical word one or more times did not lead to more false recognition than no rehearsal, for either the 2-s or the 5-s rate, as shown by the means. Similarly for remember judgments, rehearsing a critical word that was later recognized did not lead to more remember judgments than no rehearsal, for the 2-s rate or the 5-s rate, $z = 1.25$, $p > .10$. The result for the 5-s remember comparison should be treated with caution, as there were relatively few observations for many individual rehearsal frequencies in this condition and considerable variability was observed in the conditional probabilities. Even so, a general conclusion can be drawn. Unlike recall, recognition and remember judgments showed no reliable effect of rehearsal of critical words during study.

The same conclusions were reached by a second analysis that calculated conditional probabilities on an individual-participant basis. The overt-rehearsal participants were again divided into four groups of 20, depending on presentation rate and test. But to be included in this analysis, each participant had to have rehearsed and not rehearsed at least one critical word, so that two conditional probabilities could be calculated for each person: the conditional false memory probability for critical words that were rehearsed one or more times and the conditional probability for critical words that were not rehearsed. Table 5 presents these results for recall, recognition, and remember judgments for both presentation rates. The number of participants in each condition is less than 20 because not everyone in a condition yielded a pair of conditional probabilities, as rehearsal was determined entirely by each participant.

Once again, participants recalled previously rehearsed critical words with greater probability than those not rehearsed, but rehearsal had no effect on recognition or remember judgments. These results were supported by analyses of variance that treated presentation rate as a between-participants variable and rehearsal status (rehearsed vs. not rehearsed) as a within-participants variable. Rehearsed critical words were recalled more than nonrehearsed critical words, $F(1, 28) = 6.25$, $MSE = 0.51$, $p < .05$, but they were not recognized more or judged as remembered more, both Fs < 1.0, and there were no reliable interactions of rehearsal condition and presentation rate.

Thus, even though the conditional probability means for rehearsed and nonrehearsed critical words were calculated by different procedures for Tables 4 and 5, the similarity in the means and statistical outcomes attests to the reliability of our results. The answer to our second research question is thus also unambiguous: Critical-word rehearsal at study enhances subsequent false recall, but has no reliable effect on false recognition or remember judgments. Thus, false memory in the DRM task is not dependent on the prior conscious rehearsal of the critical words during study, regardless of memory measure.

These findings are consistent with the activation-monitoring approach and fuzzy-trace theory, as both explanations include a fast-acting, largely automatic activation or generation process that can activate representations for critical words or generate gist representations during study. These same results are inconsistent with explanations, such as Underwood's activation hypothesis, that hold that false memory is the result of a conscious process of covert verbalization of

Table 5. *Mean conditional probabilities of recall, recognition, and remember judgments for unrehearsed and rehearsed critical words*

Memory measure and presentation rate	Critical-word status	
	Unrehearsed	Rehearsed
Recall		
2 s (n = 15)	.19	.44
5 s (n = 15)	.16	.29
Recognition		
2 s (n = 12)	.72	.62
5 s (n = 18)	.58	.59
Remember		
2 s (n = 11)	.83	.70
5 s (n = 12)	.34	.59

Note. The n in each condition reflects the number of participants who contributed data to this analysis, which required that the participant rehearsed at least one critical word and did not rehearse at least one critical word.

critical words that leads to misattribution errors (cf. Mather et al., 1997). We observed false memory regardless of whether the critical words were consciously rehearsed or not. Conscious rehearsal may merely augment this automatic process. Moreover, explanations of incorrect remember judgments that assume that critical words need to be consciously experienced during study also face difficulty with our results. As Brainerd et al. (2001) noted, these metamemory judgments could be confabulations based on the repeated encoding of similar word meanings during study.

Why did critical-word rehearsal enhance false recall, but not false recognition or remember judgments? One possibility is that a recognition test is less sensitive than a recall test to the effects of different variables in this procedure (Roediger, McDermott, & Robinson, 1998). If so, our conclusion about the lack of effect of critical-word rehearsal on false recognition or remember judgments should be treated with caution. The fact remains, however, that critical-word rehearsal was unnecessary for false memory, regardless of how memory was measured.

This finding strongly indicates that false memory in the DRM procedure is linked to an automatic activation or generation process, and it is consistent with three other lines of converging evidence. First, false memory in this procedure is observed even with rapid presentation rates during study that lead to severe reductions in accurate memory (Roediger et al., 2002; Seamon et al., 1998). Second, when participants are forewarned of this illusion, they still yield a substantial false memory effect, suggesting that this illusion is amenable to only limited conscious control (Gallo, Roberts, & Seamon, 1997; McDermott & Roediger, 1998). Third, directed forgetting instructions reduce accurate, but not false, recall in the DRM procedure, again indicating a limit on conscious control (Kimball & Bjork, in press; Seamon et al., in press). Collectively, these findings provide compelling converging evidence that an automatic activation or generation process underlies false memory in the DRM procedure, although, as researchers have previously suggested (McDermott, 1997; Roediger et al., 2001; Seamon et al., 1998), conscious processing of the critical words may also contribute to false memory. The present research suggests that a secondary conscious process involving the rehearsal of critical words can supplement the primary automatic activation or generation process and lead to enhanced false memory, at least when it is tested by free recall.

Acknowledgments—We thank Charles Brainerd and Kathleen McDermott for helpful comments.

REFERENCES

Brainerd, C.J., & Reyna, V.F. (1996). Mere memory testing creates false memories in children. *Developmental Psychology*, *32*, 467–478.

Brainerd, C.J., & Reyna, V.F. (1998). Fuzzy trace theory and children's false memories. *Journal of Experimental Child Psychology*, *71*, 81–129.

Brainerd, C.J., Wright, R., Reyna, V.F., & Mojardin, A.H. (2001). Conjoint recognition and phantom recollection. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 307–327.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*, 17–22.

Gallo, D.A., Roberts, M.J., & Seamon, J.G. (1997). Remembering words not presented in lists: Can we avoid creating false memories? *Psychonomic Bulletin & Review*, *4*, 271–276.

Goodwin, K.A., Meissner, C.A., & Ericsson, K.A. (2001). Toward a model of false recall: Experimental manipulation of encoding context and the collection of verbal reports. *Memory & Cognition*, *29*, 806–819.

Kimball, D.R., & Bjork, R.A. (in press). The influence of intentional and unintentional forgetting on false memories. *Journal of Experimental Psychology: General*.

Koutstaal, W., & Schacter, D.L. (1997). Gist-based false recognition of pictures in older and younger adults. *Journal of Memory and Language*, *37*, 555–583.

Mather, M., Henkel, L.A., & Johnson, M.K. (1997). Evaluating characteristics of false memories: Remember/know judgments and memory characteristics questionnaire compared. *Memory & Cognition*, *25*, 826–837.

McDermott, K.B. (1997). Priming on perceptual implicit memory tests can be achieved through presentation of associates. *Psychonomic Bulletin & Review*, *4*, 582–586.

McDermott, K.B., & Roediger, H.L., III. (1998). Attempting to avoid illusory memories: Robust false recognition of associates persists under conditions of explicit warnings and immediate testing. *Journal of Memory and Language*, *39*, 508–520.

McDermott, K.B., & Watson, J.M. (2001). The rise and fall of false recall: The impact of presentation duration. *Journal of Memory and Language*, *45*, 160–176.

Roediger, H.L., III, Balota, D.A., & Robinson, K.J. (2002). *Automatic mechanisms in the arousal of false memories*. Unpublished manuscript, Washington University, St. Louis, MO.

Roediger, H.L., III, Balota, D.A., & Watson, J.M. (2001). Spreading activation and arousal of false memories. In H.L. Roediger, III, J.S. Nairne, I. Neath, & A.M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 95–115). Washington, DC: American Psychological Association.

Roediger, H.L., III, & McDermott, K.B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 803–814.

Roediger, H.L., III, & McDermott, K.B. (2000). Distortions of memory. In E. Tulving & F.I.M. Craik (Eds.), *The Oxford handbook of memory* (pp. 149–162). New York: Oxford University Press.

Roediger, H.L., III, McDermott, K.B., & Robinson, K.J. (1998). The role of associative processes in creating false memories. In M.A. Conway, S.E. Gathercole, & C. Cornoldi (Eds.), *Theories of memory* (Vol. II, pp. 187–245). Hove, England: Psychology Press.

Rundus, D. (1971). Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology*, *89*, 63–77.

Seamon, J.G., Luo, C.R., & Gallo, D.A. (1998). Creating false memories of words with or without recognition of list items: Evidence for nonconscious processes. *Psychological Science*, *9*, 20–26.

Seamon, J.G., Luo, C.R., Schlegel, S.E., Greene, S.E., & Goldenberg, A.B. (2000). False memory for categorized pictures and words: The category associates procedure for studying memory errors in children and adults. *Journal of Memory and Language*, *42*, 120–146.

Seamon, J.G., Luo, C.R., Schwartz, M.A., Jones, K.J., Lee, D.M., & Jones, S.J. (2002). Repetition can affect accurate and false recognition similarly or differently. *Journal of Memory and Language*, *46*, 323–340.

Seamon, J.G., Luo, C.R., Shulman, E.P., Toner, S.K., & Caglar, S. (in press). False memories are hard to inhibit: Differential effects of directed forgetting on accurate and false recall in the DRM procedure. *Memory*.

Snodgrass, J.G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, *117*, 34–50.

Toglia, M.P., Neuschatz, J.S., & Goodwin, K.A. (1999). Recall accuracy and illusory memories: When less is more. *Memory*, *7*, 233–256.

Underwood, B.J. (1965). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology*, *70*, 122–129.

(RECEIVED 10/29/01; REVISION ACCEPTED 1/30/02)